EMBEDDED NON-VOLATILE MEMORY

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a Continuation of co-pending U.S. patent application Ser. No. 14/733,919, filed Jun. 8, 2015, which application is a Continuation of U.S. patent application Ser. No. 14/306,801, filed Jun. 17, 2014, now U.S. Pat. No. 9,054,031, which application is a Continuation of U.S. patent application Ser. No. 13/707,895, filed Dec. 7, 2012, now U.S. Pat. No. 8,786,023, which application claims benefit of U.S. Provisional Patent Application Ser. No. 61/630,297, filed Dec. 8, 2011 and U.S. Provisional Patent Application Ser. No. 61/632,393, filed Jan. 20, 2012. Each of the aforementioned related patent applications is herein incorporated by reference.

FIELD OF THE INVENTION

[0002] The present invention relates to the manufacture and processing of semiconductors, and more particularly to the manufacture and processing of embedded memory on CMOS logic.

REFERENCE TO A SEQUENCE LISTING, A TABLE, OR A COMPUTER PROGRAM LISTING COMPACT DISK APPENDIX

[0003] Not Applicable.

REFERENCE REGARDING FEDERAL SPONSORSHIP

[0004] Not Applicable.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0005] Not Applicable.

SUMMARY OF THE INVENTION

[0006] Microcomputer chips and microcontroller chips typically have non-volatile memory for holding executable code and volatile memory for holding data and register values. More and more, microcontroller chips are providing small areas of non-volatile memory for holding key data values. For example, the application of a microcontroller would benefit if setup parameters could be stored in non-volatile memory along with executable code so that operation could resume after power is restored without having to reengage the setup process.

[0007] Presently, CMOS microcontroller chips incorporate small arrays of Flash memory for retention of key data values. Many CMOS microcontrollers also put executable code in Flash memory as well. However, in addition to CMOS' processing masks and processing steps, Flash memory requires several additional masks and their associated processing steps, whereby many of these additional masks are at the critical geometry of the overall process and therefore of the most costly class of mask. Furthermore, the additional processing steps for incorporating this non-volatile memory must not compromise the existing CMOS process nor significantly change that process. What is needed is a non-volatile memory that can be incorporated into a CMOS process that requires few additional masks and processing steps.

[0008] The present invention is a method of incorporating a non-volatile memory into a CMOS process that requires four or fewer masks and limited additional processing steps.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] FIG. **1-76** are illustrations of a semiconductor substrate, following a sequence of processing steps, the result of each processing step being depicted in cross section.

[0010] FIG. **77** illustrates a semiconductor substrate, following the application of multiple metal layer steps, the result being depicted in cross section.

DETAILED DESCRIPTION

[0011] Today's CMOS processes are finely tuned, complex processes that are very sensitive to being broken if changes are made. One of the largest manufacturers of CMOS logic, the Intel Corporation, addresses this sensitivity with its slogan, "Copy exactly." In other words, even the slightest of changes can disrupt the entire process and require a re-tuning of that process. Various individual process steps and sequences are likewise very sensitive to being disrupted.

[0012] Some of the most delicate parts of the process include the silicon surface preparation of the areas of the substrate where the channels of MOS transistors are going to be put. These surfaces must be kept near perfect to avoid current leakage paths and recombination sites that would degrade performance.

[0013] Another sensitive stage is following the formation of salicided contacts to the MOS transistors as these contacts are easily damaged after formation by elevated temperatures (e.g., a temperature of 800.degree. C. or higher).

[0014] The addition of a non-volatile memory array benefits if that array can be very compact. Typically, the most compact memory arrays have a 4F.sup.2 memory cell size. Typically a 4F.sup.2 memory array is achieved with a diode matrix wherein the diodes are formed in the vertical orientation. Furthermore, these diodes must have low leakage to prevent unnecessary power loss. Ideally, an epi-silicon diode formation process will yield the highest quality diodes. However, epi-silicon is grown at elevated temperatures.

[0015] Formation of one or more diode arrays within a CMOS process is problematic. If an array is formed after the CMOS process is completed, the thermal budget of the diode array processing will damage the salicided contacts of the MOS transistors. If a diode array is formed prior to the CMOS process, the delicate surface areas of the wafer where the MOS transistor channels are to be formed can be compromised. The present invention is a means of incorporating an epi-silicon (or poly-silicon), non-volatile diode memory array within a CMOS process without compromising the performance or characteristics of the resulting CMOS logic. The present invention is an epi-silicon process sequence that is introduced into a standard CMOS process (i) after the MOS transistors' gate oxide is formed and the gate poly-silicon is deposited (thereby protecting the delicate surface areas of the MOS transistors) and (ii) before the salicided contacts to those MOS transistors are formed (thereby performing any newly introduced steps having an elevated temperature, such as any epi-silicon or poly-silicon deposition for the formation of diodes, prior to the formation of that salicide). This timing of the diode formation steps has